Goodness-of-Fit Testing for Hawkes Processes

Reproduction and Application of an Asymptotically Distribution-Free Test

Adrien Damez Arnaud Van de Velde Elena Perotti
Angelo Picerno Gavriil Kharkhodine

Project P15 — Hawkes Processes

January 2026

Motivation

- Hawkes processes are standard for **self-exciting event data**
- Widely used in:
  - High-frequency finance
  - Seismology, epidemiology, social networks
- Key question: how to assess whether a fitted Hawkes model is adequate?

Classical Goodness-of-Fit

- Standard approach: **Random Time Change (RTC)**

$$\tau_i = \int_0^{t_i} \lambda_{\hat{\theta}}(s)\, ds$$

- Under the true model: transformed inter-arrival times $\sim \mathrm{Exp}(1)$
- In practice:
    - Parameters unknown
    - Plug-in estimation error ignored
- Consequence: tests may be badly sized or overly sensitive

Main Idea of the Project

Reference

Baars, Can & Laeven (2025)

- Propose an **asymptotically distribution-free** GOF test
- Explicitly accounts for:
  - Parameter estimation uncertainty
- Based on a **martingale transformation**

Our goals:

- Reproduce their results for Hawkes processes
- Compare with naive and RTC-based procedures
- Apply to real high-frequency financial data

Hawkes Process Model

$$\lambda_\theta(t) = \mu + \sum_{t_i < t} g_\theta(t - t_i)$$

- $\mu > 0$: baseline intensity (constant)
- g_θ: excitation kernel
- Stationarity condition:

$$\int_0^\infty g_\theta(s)\, ds < 1$$

Kernels considered:

- Exponential
- Power-law
- Multi-exponential

Parameter Estimation

- Parameters estimated by **Maximum Likelihood**

$$\ell_T(\theta) = \sum_{i=1}^{N(T)} \log \lambda_\theta(t_i) - \int_0^T \lambda_\theta(t)\, dt$$

- Efficient recursion for exponential / multi-exponential kernels
- Power-law kernel: higher computational cost

Important: treating $\hat{\theta}$ as fixed can distort GOF tests

Goodness-of-Fit Procedures Compared

1. **Naive compensated process**
2. **Transformation-based procedure** (Baars et al.)
3. **Naive Random Time Change (RTC)**

All tests use classical statistics:

- Kolmogorov–Smirnov
- Cramér–von Mises
- Anderson–Darling

Transformation-Based Test (Key Idea)

- Start from the compensated empirical process

$$\eta_T(u) = \frac{1}{\sqrt{T}} \left(N(uT) - \int_0^{uT} \lambda_{\hat{\theta}}(s)\, ds \right)$$

- Apply a transformation that removes estimation effects
- Obtain increments that are asymptotically i.i.d. $\mathcal{N}(0, 1)$

Key advantage

Asymptotic distribution does **not depend** on θ

Simulation Study

- Hawkes processes simulated under the null
- Same setup as Baars et al. (2025)
- $T = 5000$, 500 Monte Carlo replications
- Compare rejection frequencies at 1%, 5%, 20%

Models tested:

- Correct specification
- Misspecified kernels

Simulation Results — Main Message

- **Transformation-based test**
  - Correctly sized under the null
  - Robust across kernel families
- **Naive test**
  - Strongly conservative
- **Naive RTC**
  - Undersized under the null
  - Rejects aggressively under misspecification

Application to Real Data

- High-frequency trade data (Société Générale)
- 12 trading days, 9:00–17:30
- Data split into **one-hour windows**
- Baseline intensity assumed locally constant

Model used: multi-exponential Hawkes

Estimation Strategy on Real Data

- Full MLE unstable for multi-exponential kernels
- Adopt a **two-step approach**:
  1. Estimate decay parameters via GMM on inter-arrival times
  2. Fix decay rates and estimate remaining parameters
- Improves numerical stability

Empirical GOF Results

- Transformation-based test:
  - Rejects more than naive test
  - Provides a sensitive diagnostic
- Naive RTC:
  - Rejects almost systematically
  - Extremely sensitive to plug-in estimation

Why Does RTC Fail?

- Q–Q plots of RTC-transformed inter-arrival times
- On simulated data: close to Exp(1)
- On real data:
  - Global deviations from exponentiality
  - Not only tail effects

Interpretation

RTC amplifies mild misspecification + estimation error

Conclusion

- Accounting for parameter estimation is **crucial**
- Transformation-based GOF:
  - Correctly sized
  - Interpretable on real data
- Naive RTC can be misleading in practice

Takeaway: GOF tests should be used as diagnostic tools, not mechanical decisions

Perspectives

- Data-driven choice of (n, τ)
- Time-varying baseline intensities
- Multivariate Hawkes processes
- Estimation-aware alternatives to RTC